

# Advance demand information and a restricted production capacity: on the optimality of order base-stock policies

Jacob Wijngaard · Fikri Karaesmen

Published online: 13 January 2007  
© Springer-Verlag 2007

**Abstract** This paper considers the optimality of order aggregation in a single-item production/inventory problem with advance demand information and a restricted production capacity. The advance demand information is modeled by introducing a positive customer order lead time. The paper proves, when customer order lead times are less than a threshold value, it is allowed to aggregate the orders over time when establishing the optimal production decision. This implies the optimality of an order base-stock policy. It shows also that in case of linear inventory cost, the positive effect of advance demand information is equal to a cost reduction that is proportional to idle time and foreknowledge horizon. The results hold for the backlogging case as well as for the lost-sales case.

**Keywords** Advance demand information · Restricted production capacity · Base-stock policies · Backlogging and lost-sales · Sample path equivalence · Optimality and heuristics

## 1 Introduction

The standard modeling assumption in the analysis of stochastic production/inventory systems implies that the random customer demand (whose probability

---

J. Wijngaard (✉)  
Faculty of Management and Organization, University of Groningen, P.O. Box 800,  
9700 AV Groningen, The Netherlands  
e-mail: j.wijngaard@bdk.rug.nl

F. Karaesmen  
Department of Industrial Engineering, Koc University, 34450 Sariyer, Istanbul, Turkey  
e-mail: fkaraesmen@ku.edu.tr

distribution is known) has to be satisfied immediately. If that is not possible, the system incurs a penalty such as a backordering cost proportional to the delay or the forgone revenue due to the lost-sale. In several retail or industrial settings, however, this assumption does not apply, since there may be advance demand information (ADI) on the customer demand such as early customer orders with specified due dates. A relatively recent stream of research extends the standard modeling framework to cases where there is ADI.

The main difficulty in the analysis of a stochastic model that encompasses ADI is that the state description of the production/inventory system must include the arrival times and due dates of each order in addition to the current inventory position. In general, the optimal production control (or replenishment) policy is complicated under ADI, simply because it depends on the full state information. This is in contrast with production/inventory systems without ADI for which *base-stock policies* are known to be optimal under general conditions in the absence of setup costs. The optimality holds for the discounted cost/reward case as well as for the average cost/reward case. It results from the fact that there is an ideal inventory one wants to start with each period. In case of continuous review systems, one wants to go back to this ideal inventory as soon as there has been a demand (see Beckmann 1961). If the production lead time is 0, this optimality holds for backlogging as well as lost-sales and for a broad range of inventory and backlogging costs. It depends only on the assumption of system stationarity (demand and cost functions) and on the absence of setup costs. If the production lead time is positive, the optimality still holds for the backlogging case; instead of looking at the costs from now on, one looks at the costs beyond the production lead time (see Karlin and Scarf 1958). The optimality holds also in case of capacitated models where the processor can be turned off or on, for backlogging as well as for lost-sales (see Zipkin 2000). We will use this type of processor control also in this paper.

In addition to optimality, the base-stock policy is very attractive due to its simplicity: it releases a production order for each demand removal. This simplicity (and optimality) suggests using a similar logic for a production/inventory system with ADI. Instead of the inventory itself, one may use the inventory minus the order book (i.e., all future demand that is already ordered). This is an *order base-stock policy* (see Hariharan and Zipkin 1995). The order base-stock policy is known to be optimal in a number of special cases that comprise ADI. One example is the continuous review inventory system with Poisson demand and exogenous replenishments (lead time  $\ell$ ) with constant customer lead times (denoted by  $h$ ), analyzed by Hariharan and Zipkin (1995). This result follows from the combination of two other results. In the first place that in case of an un-capacitated system, ADI beyond the replenishment lead time is useless. In the second place that ADI over a horizon  $h < \ell$  makes the system equivalent to the system without ADI and production lead time  $\ell - h$ . In that case the inventory plus work-in-process minus the demand as far as it is known responds similarly to production and demand as the inventory plus work-in-process does in the system without ADI. For a periodic review system with exogenous replenishments and more complex ADI, Gallego and Ozer (2001)

obtain a similar result. They introduce the term *modified inventory* for inventory plus work-in-process minus demand as far as it is known. They also show that the optimal policy is more complicated and is shown to depend on the complete advance demand vector if the setup cost is positive.

For capacitated (endogenous replenishment) systems, the situation is already complicated without setup costs. In this case, the replenishment lead time is no longer a constant since it depends on the level of congestion at the production facility. Karaesmen et al. (2002) consider the discrete version of the M/M/1 make-to-stock queue; the customer orders are assumed to have the same order lead time. Based on numerical observations, they note that an order base-stock policy is optimal if the customer order lead time is less than a given threshold. In addition, they propose a reasonable control policy in case of long customer order lead times, where future orders beyond the critical threshold are not taken into account. The purpose of this paper is to provide a complete proof of the result that an order base-stock policy is optimal when the customer order lead time is less than a threshold value which can be characterized. The production resource is assumed to be reliable and production is continuous (see Wijngaard 2004). The backlogging case and the lost-sales case are combined. If the realized utilization rate of the system is denoted by  $\rho$ , the customer lead time is  $h$  and  $P$  is the production speed, the optimality of order base-stock policies implies also that the effect of ADI is a reduction of inventory by  $(1 - \rho)hP$ . Combining backlogging and lost-sales leads also to a new heuristic in case of a long customer order lead time that is an adapted order base-stock policy. The approach is based on the construction of equivalence of the system with and the system without ADI. The paper of Wijngaard (2004) touches on this issue already. But here the equivalence is worked out more completely and precisely. It appears to facilitate a unified proof of the optimality of order base-stock policies for backlogging and lost-sales and for combinations.

Section 2 presents the literature review. Section 3 describes the model. Section 4 gives the main results for the case with a customer order lead time that is smaller than a certain threshold value. Section 5 gives extensions and limitations of these results. Section 6 discusses heuristics for the case with a longer customer order lead time. Section 7 presents conclusions and suggestions for further research.

## 2 Literature review

The literature on inventory models with ADI is rapidly growing. The literature review focuses on papers that are related to the optimality of order base-stock policies. A number of papers investigate the uncapacitated replenishment situation (i.e., exogenous replenishment lead times) in case of ADI. Since our focus is on the capacitated system, we only briefly review two of the main related results for uncapacitated systems. Hariharan and Zipkin (1995) analyze the standard continuous review inventory system with Poisson demand arrivals, constant replenishment lead times, and constant customer lead times. They show that an

order base-stock policy is optimal whenever the customer lead time is smaller than the replenishment lead time and characterize the reduction in inventory related costs as a function of the customer lead time. Gallego and Ozer (2001) investigate a periodic review system with exogenous replenishments and a more complicated ADI structure where customer lead times are not constant. They also obtain a similar result when the customer order information horizon is smaller than the replenishment lead time; an order base-stock policy is optimal then.

There is also another stream of research which at first is not directly related with ADI but has certain similarities from a modeling perspective. Simpson (1958) proposes an approach to compute the optimal safety stock levels in multi-echelon systems. Interestingly, each echelon in his model has an incoming lead time (corresponding to a supply lead time) and an outgoing lead time (corresponding to the customer lead time). Inderfurth (1991) extends the analysis of Simpson to more general supply-chain structures. Most of these models treat uncapacitated multi-echelon systems and assume that demand that cannot be satisfied within the outgoing lead time is sourced from alternative sources. Graves and Willems (2003) provide a review of this literature.

For capacitated systems under continuous review which is the main focus of this paper, Buzacott and Shanthikumar (1994) consider an M/M/1 make-to-stock queue and investigate the effects of ADI with constant customer lead times. They analyze order base-stock policies and characterize the inventory cost reduction. The main difficulty in addressing optimality issues in continuous time is that the system state information must include all future order times in addition to the inventory position. In order to investigate optimality issues, Karaesmen et al. (2002) use the discrete approximation of the M/M/1 make-to-stock queue by modeling order inter-arrival times and processing times as geometric distributions. For this model, they show that generalized base-stock policies are optimal. These policies are relatively complicated since they not only include inventory position but also future order arrival times in the replenishment decisions. On the other hand, it is numerically found that order base-stock policies (i.e., a subclass of generalized base-stock policies) are optimal when the customer lead time is smaller than a threshold value. Moreover, this threshold value is a function of the replenishment lead time distribution. Karaesmen et al. (2003, 2004) investigate further properties of order base-stock policies in M/M/1 and M/G/1 type make-to-stock queues, using analytical results and approximations. Liberopoulos and Koukourmialos (2005) investigate both single-stage and two-stage systems using simulation and present additional results. Finally, a paper by Gayon et al. (2004) considers a multi-class problem and investigates stock rationing issues in the context of ADI. This paper addresses optimality issues by using exponentially distributed customer lead times which facilitates the state representation significantly.

A different stream of research considers capacitated production/inventory problems in discrete time. Gullu (1995, 1996) models ADI coming from an external forecasting process using the Martingale Model of Forecast Evolution. The optimal base-stock level is shown to depend on the forecast vector.

Toktay and Wein (2001) further analyze this model using a heavy-traffic approximation and present approximations for the optimal base-stock level. Ozer and Wei (2004) use a different representation of ADI that only allows additive updates. In general, the optimal base-stock level for this model also depends on the future demand information. However, when the information horizon is smaller than the production lead time, an order base-stock policy is optimal.

The contribution of this paper with respect to the above literature is as follows. First, a number of the above papers (Toktay and Wein 2001; Karaesmen et al. 2003, 2004; Liberopoulos and Koukourmialos 2005) use order base-stock policies as effective heuristics. Numerical results in Karaesmen et al. (2002) suggest that such policies are optimal if the customer lead time is smaller than a given threshold, but a formal proof is lacking so far. The proof of this optimality along with a characterization of the critical threshold is provided here for the M/D/1 make-to-stock queue. Second, this proof holds not only for the backorder case but also for the case of lost-sales (order rejection). This combination of lost-sales and backlogging also leads to a new heuristic for the backlogging case with long customer lead time. Third, some of the above papers (Toktay and Wein 2001; Karaesmen et al. 2003, 2004) provide an approximate characterization of the cost reduction as a function of the customer lead time. An exact characterization of this reduction is provided here. This paper builds on partial results in Wijnngaard (2004). The proofs in this paper are more general and show more completely the scope of the approach.

### 3 Model

The model is a single-product, make-to-stock model. Demand is compound Poisson, with arrival rate  $\lambda$ . The customer order size is integer, with distribution function  $F(\cdot)$ . The required customer lead time is equal to  $h$ . So, the orders are known  $h$  time units in advance (the foreknowledge horizon). This is a standard assumption if the orders come from a downstream supply chain member using an MRP-type replenishment policy (see Buzacott and Shanthikumar 1994; Karaesmen et al. 2002). The case  $h = 0$  corresponds to the system without ADI. There is an order acceptance mechanism that determines whether orders are accepted or rejected. Once accepted, orders are placed in the order book. On the due date, the customer order is compared with the available stock. If there is sufficient stock, the order is delivered. Otherwise, the order backlog is increased by the customer order size. The server produces at rate 1 (the production rate  $P$  may be normalized to 1, without loss of generality, by changing the unit of time). Production is reliable and continuous. After a unit of product is completed, the server can be turned off or it can continue producing. The starting inventory is assumed to be integer. Since production runs are equal to 1 and order sizes are integer, the non-integer part of the inventory, at those points in time where a new production run may be started, is equal to the non-integer part of the starting inventory. This makes the assumption of an integer starting inventory natural. The system objective is to maximize the average reward per unit of

time. The reward function is the difference between net sales revenues and cost for keeping inventory plus cost of backlogging. The inventory cost is assumed to be linear:  $c$  per unit per unit of time. The cost of backlogging,  $b(\cdot)$ , is assumed to be an increasing function of the size of the backlog. The net sales revenue (difference of revenue and direct cost per unit) is assumed to be proportional to the order size:  $r$  per unit. The unit of measure of  $c$ ,  $b(\cdot)$ , and  $r$  is a cost measure, like € or \$. So the objective function is

$$\max_p \lim_{T \rightarrow \infty} \frac{1}{T} E_p \left[ rW(T) - \int_0^T (cI^+(t) + b(I^-(t))) dt \right],$$

where  $p$  denotes the policy,  $W(t)$  denotes the number of orders delivered until time  $t$ ,  $I^+(t) = \max(I(t), 0)$  and  $I^-(t) = \max(-I(t), 0)$ . If backlogging is not allowed, then  $I^-(t) = 0$ . If the limit does not exist, the  $\liminf$  is taken (see Ross 1970; Puterman 1994).

To model and evaluate the decision making, the system is observed at the following points in time:

- Production run finishes.
- Order arrivals (just after arrival, before the acceptance decision).

When a production run ends on a certain point  $t$ , a plan is made on when to start the next run ( $t_s \geq t$ ). It is also allowed to plan not to start a new run ( $t_s = \infty$ ). If no order arrives between  $t$  and  $t_s$ , the next run is started indeed at  $t_s$ . If there is an earlier order arrival, this arrival time is the next observation point and the plan for the next production start is updated. Orders arrive during production and during idle periods. When an order arrives during a production run, the only decision is the order acceptance decision. It is not necessary then to make a new plan on the start of the next run. This plan will be updated again when the production run finishes. When an order arrives during an idle period, it is also necessary to update the planned next production start. The state of the system at an observation point in time  $t$  can be represented by  $\{X(t), t \geq 0\}$ , with  $X(t) = (o(t), q(t), I(t), \text{OB}(t), l(t))$ , where

- $o(t)$  boolean indicating the type of the observation point (arrival or run finish),
- $q(t)$  size of the arriving customer order,
- $I(t)$  inventory at time  $t$ ,
- $\text{OB}(t)$  the order book at time  $t$ ,
- $l(t)$  the remaining length of the actual production run at time  $t$ .

The variable  $q(t)$  is only relevant in case of an order arrival. In case of a production run finish, we define  $q(t) = 0$ . The remaining length of the actual production run is only positive at an observation point, if an order arrives during a production run. It determines the first possibility to start a new production run. Orders are numbered in sequence of arrival. Each order can be characterized by its due date ( $dd_i$ ) and its size ( $q_i$ ). For the state of the system at time  $t$ , only

the orders in the order book are relevant. These are the orders that have been accepted and not delivered yet. To characterize the order book we do not use the due dates, but the relative due dates ( $rd_i = dd_i - t$ ):

$$OB(t) = \{(rd_i, q_i), i | i \text{ the order numbers of the orders that are available on time } t\}$$

The development of  $I(t)$ ,  $OB(t)$ , and  $I(t)$  until the next observation point follows deterministically from the starting observation point. The relative due dates in the order book are decreasing constantly at speed 1. In due time, an order is delivered and removed from the order book and the inventory drops with an amount that equals the order size. During production, the inventory is growing at speed 1. The state of the system at the next observation point is the result of this deterministic process, possibly cut off by an order arrival that comes before the planned next production start.

The process with these observation points, state definition and action possibilities is a semi-Markov decision process (Ross 1970; Puterman 1994). A policy prescribes an allowable action on each of the observation points. The optimal policy maximizes the average reward per unit time. This semi-Markov formulation is going to be used to explore the equivalence of the system with ADI and the system without ADI. This equivalence is going to be the basis for the results.

The *virtual inventory* helps to reveal this equivalence. The virtual inventory is defined as

$$V(t) = I(t) - C(t) + 1 \cdot h, \quad (*)$$

with  $C(t)$  being the sum of all customer order sizes in the order book  $OB(t)$ , and thus the time aggregate of the customer orders. That means that the virtual inventory  $V(t)$  is the inventory that results after  $h$  time units (at  $t + h$ ) if there is no idle time until  $t + h$  (recall that the production is either on or off and that the speed is equal to 1). The virtual inventory determines whether it is feasible to accept a new customer order without a stock-out; if  $V(t) \geq z$ , it is possible to accept a customer order of size  $z$ . The behavior of  $V(t)$  in the system with ADI is also comparable to the behavior of  $I(t)$  in the system without ADI: production leads to a steady increase, acceptance of an order leads to an immediate drop. There is a *difference* in performance between both systems, however. Let  $I'(t)$  denote the inventory level at time  $t$  in the system without ADI. In this system, the performance (with respect to inventory and backorders) is based on  $I'(t)$ . In the system with ADI, the inventory cost is not based on  $V(t)$ , the state variable that corresponds to the state variable  $I'(t)$  in case of no ADI, but on  $I(t)$ , and  $I(t)$  is on average lower than  $V(t)$ . Since we are interested in the long-run average reward, we may just as well look at  $I(t + h)$  instead of at  $I(t)$ . And there is a rather simple relationship of  $V(t)$  and  $I(t + h)$ :

$$I(t + h) = V(t) - L(t, t + h) \cdot 1, \quad (**)$$



with  $I(t+h)$  being the inventory at  $t+h$  and  $L(t, t+h)$  the idle time during the interval  $(t, t+h)$ . So,  $L(t, t+h) \cdot 1$  is the lost production. In the sequel (next section) we are going to use the fact that under an order base-stock policy this idle time between  $t$  and  $t+h$  depends only on  $V(t)$  and the order arrivals beyond  $t$ .

The relationship of both systems becomes even clearer if we replace the order book  $OB(t)$  in the state description by the set of drops of  $V(\cdot)$  during the interval  $(t-h, t)$ . The positions of the drops are the arrival times of the orders that arrived during  $(t-h, t)$ , the size of the drops refers to the size of the corresponding orders. The set of drops is denoted by  $H_v(t)$ . To realize a state description in  $V(\cdot)$ , we also replace  $I(t)$  by  $V(t)$ . The relationship (\*) allows this. The variables  $V(t)$  and  $H_v(t)$  are completely equivalent to the variables  $I(t)$  and  $OB(t)$ . Through  $I(t)$  and  $OB(t)$  we observe the inventory at time  $t$  and look ahead to the due dates of the orders in the order book, through  $V(t)$  and  $H_v(t)$  we observe the virtual inventory and look back at the arrival dates of the orders that have not been delivered yet. Let  $Y(t)$  be the state formulation using  $V(t)$  and  $H_v(t)$ :

$$Y(t) = (o(t), q(t), V(t), H_v(t), l(t)).$$

Base-stock policies are of particular importance for this paper. With a base-stock policy for the system without ADI, production continues as long as  $I(t)$  is less than a target level  $S$ , which is called the base-stock level; production is shut down if  $I(t)$  reaches  $S$ . Similarly in case of ADI, with an order base-stock policy, production continues as long as the virtual inventory  $V(t)$  is below the base-stock level.

Note that the base-stock character is only associated with the production part of the policy, not with the acceptance part. If all customer orders are size 1, it is optimal in case of no ADI to accept orders if  $I(t) \geq 1$ . This transfers to  $V(t) \geq 1$  in case of ADI. In case of unequal order sizes, optimal acceptance may be more complicated.

The above state representation makes it possible to translate policies for the system with ADI into policies for the system without ADI. The corresponding patterns of  $V(t)$  and  $I(t)$  are identical. This result is formulated in Lemma 1. The difference of both systems is only in the objective function.

**Lemma 1** *An arbitrary policy  $p^h$  for the system with ADI, with state representation  $Y(\cdot)$ , can be converted to a policy  $p^0$  for the system without ADI by replacing the dependence on  $V(\cdot)$  by the dependence on  $I(\cdot)$ . Suppose the system with and without ADI start with an empty order book and the same inventory, then the behavior of the system with ADI, in terms of  $V(\cdot)$ , under policy  $p^h$ , is identical to the behavior of the system without ADI, in terms of  $I(\cdot)$ , under policy  $p^0$ .*

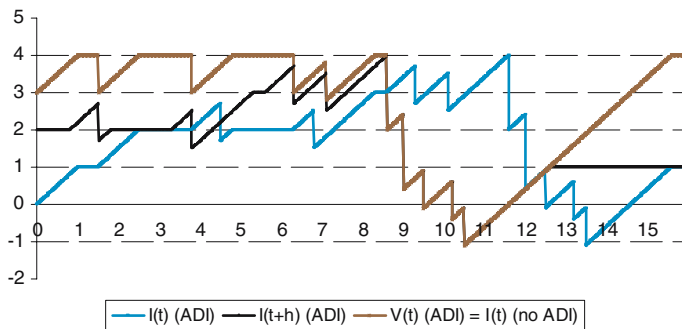
*Proof* An arbitrary policy  $p^h$  for the system with ADI [state representation  $Y(\cdot)$ ] is a function of the system history in terms of  $o(\cdot)$ ,  $q(\cdot)$ ,  $V(\cdot)$ , and  $l(\cdot)$  to the action space. This policy, with  $V(\cdot)$  replaced by  $I(\cdot)$ , can also be applied to the system without ADI. Since  $V(\cdot)$  and  $I(\cdot)$  react in the same way on production



and order acceptances, the resulting sample paths of  $V(\cdot)$  and  $I(\cdot)$  are identical if the starting states are identical.  $\square$

It is also possible, of course, to convert an arbitrary policy for the system without ADI to a policy for the system with ADI, by replacing the dependence on  $I(\cdot)$  by the dependence on  $V(\cdot)$ . Only the total content of the order book is taken into account in this way. A base-stock policy in case of no ADI leads to an order base-stock policy in case of ADI.

The result of Lemma 1 couples the behavior of  $V(t)$  in case of ADI with the behavior of  $I(t)$  in case of *no* ADI. The behavior of  $I(t+h)$  in case of ADI ( $h > 0$ ) is linked to the behavior of  $V(t)$  through equation (\*\*).



**Fig. 1** Patterns of  $V(t)$ ,  $I(t+h)$ , and  $I(t)$  for a base-stock policy with target level 4,  $h = 3$

Figure 1 gives an illustration of the behavior of  $V(t)$ ,  $I(t+h)$ , and  $I(t)$  for  $h = 3$  in case of an order base-stock policy with target level 4. The order book at  $t = 0$  is empty and  $I(0) = 0$ . So,  $V(0) = 3$ . The arrival times [sizes] of the arriving orders are 1.5 [1], 3.8 [1], 6.3 [1], 7.1 [1], 8.6 [2], 9 [2], 9.5 [1], 10.2 [1], and 10.5 [1]. So, the due dates are 4.5, 6.8, 9.3, 10.1, 11.6, 12, 12.5, 13.2, and 13.5. Since  $V(0) = 3 < 4$ , the system starts with production. The first idle period starts at  $t = 1$  when  $V(t)$  reaches 4. The system is idle on the intervals  $[1, 1.5]$ ,  $[2.5, 3.8]$ ,  $[4.8, 6.3]$ ,  $[8.3, 8.6]$ , and from 15.6 on. So  $L(0, 0+3) = 1$  and  $I(0+3) = 3 - 1 = 2$ . Because of this idle time,  $I(t+3) < V(t)$  for  $t < 8.6$ . Between 8.6 and 15.6 there is no idle time and  $I(8.6+3) = V(8.6) = 4$ . From 8.6 until 12.6,  $V(t)$  and  $I(t+3)$  are equal. Between 8.6 and 10.5 the demand is relatively high, leading to a negative  $V(t)$  on the interval  $[10.2, 11.6]$ . The pattern of  $V(t)$  as given in the figure is identical to the pattern of  $I(t)$  in case of no ADI and the application of a base-stock policy with target level 4. The inventory  $I(t+3)$  is only lower than the virtual inventory  $V(t)$  for values of  $t$  whenever both are positive; for values of  $t$  where  $V(t) < 0$ , both are equal. In the sequel we will prove that this is true in general. Since  $I(t)$  is only a translation of  $I(t+3)$  the system without ADI and the system with ADI have the same backorder performance. The advantage of ADI is the reduction of the (positive) inventory [the difference between  $V(t)$  and  $I(t+3)$ ] on the interval  $[0, 8.6]$  and from 12.6 on].

The result of Lemma 1 is used in the next sections to investigate the effect of ADI. The approach is to use the optimal policy for the system without ADI to generate an optimal (Sect. 4) or at least a good policy (Sect. 6) for the system with ADI.

#### 4 Results if customer lead times are below a certain threshold value

According to the results of the previous section, the policy for the system with ADI that corresponds to a base-stock policy for the system without ADI is an order base-stock policy. In a base-stock policy the production is continued or turned on if  $I(t) < S$ , with  $S$  the base-stock level. In an order base-stock policy, the production is continued or turned on if  $V(t) < S$ . This section shows that the policy for the system with ADI that corresponds to the optimal base-stock policy is also optimal, if  $h$  is not larger than the base-stock level.

Lemma 2 relates the average inventory for the system with ADI to the average inventory for the system without ADI.

**Lemma 2** *Let  $p^0$  be an arbitrary stationary policy for the system without ADI and let  $p^h$  be the corresponding policy for the system with ADI =  $h$ . Let  $\rho^h$  and  $\rho^0$  be the resulting utilization rates. Then  $\rho^h = \rho^0$ . Let  $\rho = \rho^h = \rho^0$ . The average inventory under policy  $p^h$  (system with  $h > 0$ ) is  $(1 - \rho)h$  smaller than the average inventory under policy  $p^0$  (system with  $h = 0$ ).*

*Proof* The behavior of  $V(\cdot)$  in the system with ADI is identical to the behavior of  $I(\cdot)$  in the system without ADI, if a proper starting state is chosen (e.g., empty order book, same inventory, see Lemma 1). Since the utilization rate does not depend on the starting state, the utilization rates in both systems are indeed equal. The difference of the average inventories follows from

$$I^h(t+h) = V(t) - L^h(t, t+h) \cdot 1,$$

with  $I^h(t+h)$  being the inventory at  $t+h$  and  $L^h(t, t+h)$  the idle time during the interval  $(t, t+h)$  of the system with  $h > 0$ , under policy  $p^h$ . By definition we have

$$E(L^h(t, t+h)) = (1 - \rho)h.$$

This completes the proof.  $\square$

The lemma says that the average inventory in the corresponding system with ADI is  $(1 - \rho)h$  lower than in the system without ADI. This inventory reduction is most attractive if it does not lead to higher backorder levels. In the example in Fig. 1, it can be observed that  $I(t+3) = V(t)$  when  $V(t) < 0$  and the inventory reduction is indeed concentrated on stretches of the time axis where the inventory is positive. If it cannot be guaranteed that the inventory reduction does not lead to higher backorder levels, it is only possible to derive a bound for the inventory cost reduction. This is consolidated in Corollary 1.

**Corollary 1** Let  $p^0$  and  $p^h$  be defined as in Lemma 2. Let  $C^0$  and  $C^h$  be the resulting average inventory cost. Then  $C^h \geq C^0 - c(1 - \rho)h$ .

*Proof* The proof follows right away from the proof of Lemma 2 and the observation that the long-run average inventory cost and backorder cost under  $p^h$  may be based on  $I^h(t + h)$  instead of  $I^h(t)$ .  $\square$

Lemma 4, Corollary 2, and Lemma 5 consider base-stock policies and (corresponding) order base-stock policies. For such policies it is easy to see the consequences of ADI for the backorder level. In Lemma 3 it is shown that, in looking for the optimal policy for the system without ADI, it is sufficient to restrict attention to base-stock policies.

**Lemma 3** Consider the system without ADI. There is an optimal policy that is base-stock.

*Proof* For semi-Markov decision processes under very general conditions there is an optimal policy that is stationary. These conditions are certainly met for the system without ADI (see Puterman 1994). Now consider an optimal stationary policy  $p^*$ . The character of the cost function implies that  $p^*$  does not ask for production if the inventory is very high and that it does ask for production if the inventory is very low. That means that there is a smallest inventory level  $I_m$  that does not ask for a production decision under policy  $p^*$ . Starting from an (integer) inventory level  $\leq I_m$ , production is continued, until at some production finish, level  $I_m$  appears to be reached (production is one by one). The production is interrupted there and the inventory levels  $> I_m$  will never be reached. That implies that the behaviour under policy  $p^*$  is equal to the behaviour under the base-stock policy with base-stock level  $I_m$ . Starting from an inventory  $I > I_m$ , the inventory will eventually fall below  $I_m$  and stay there. From then on the behaviour is again identical to the behaviour under the base-stock policy with base-stock level  $I_m$ . That implies that the average reward under policy  $p^*$  is equal to the average reward under the base-stock policy with base-stock level  $I_m$ .  $\square$

**Lemma 4** Consider a base-stock policy for the system without ADI, with base-stock level  $S$ . The behavior that results from application of the corresponding order base-stock policy in the system with ADI has the following property:

$$L^h(t, t + S) = 0 \quad \text{if } V(t) \leq 0 \quad \text{and} \quad L^h(t, t + S) \leq V(t)/1 \quad \text{if } 0 \leq V(t) \leq S.$$

*Proof* The proof follows immediately from the fact that if  $V(t) \leq S$  at a decision time  $t$ , it lasts at least  $(S - V(t))/1$  time units before the production is turned off (longer if new orders arrive in between).  $\square$

In discussing the example of Fig. 1, we stressed already that  $I(t + 3) = V(t)$  if  $V(t) < 0$  and that  $I(t + 3) \geq 0$  if  $V(t) \geq 0$ . Lemma 4 proves that this is true in general. The difference in performance is purely due to inventory reduction. The total inventory cost reduction due to ADI is proportional to the surface of the area between  $V(t)$  and  $I(t + h)$ . In the long run that surface is

about equal  $(1 - \rho)hT$ , with  $T$  the total runtime. This leads to a reduction in the average inventory cost of  $c(1 - \rho)h$ . This is formulated in Corollary 2.

**Corollary 2** *Let  $S$  be the base-stock level of the optimal base-stock policy for the system without ADI. The order base-stock policy for the system with ADI =  $h$  ( $h \leq S$ ) that corresponds to this optimal policy, has the same utilization rate (and average net revenue), the same average stock-out cost, and its inventory cost is  $c(1 - \rho)h$  lower.*

*Proof* From Lemma 1 it follows that the utilization rate is the same for the system with and the system without ADI (see also Lemma 2). Lemma 4 shows  $V(t) \leq 0 \Rightarrow L^h(t, t+S) \cdot 1 = 0$  and  $V(t) \geq 0 \Rightarrow V(t) - L^h(t, t+S) \cdot 1 \geq 0$ . Because  $h \leq S$ , we have  $L^h(t, t+h) \leq L^h(t, t+S)$ . Since  $I^h(t+h) = V(t) - L^h(t, t+h) \cdot 1$ , this means that  $V(t) \leq 0 \Rightarrow I^h(t+h) = V(t)$  and  $V(t) \geq 0 \Rightarrow I^h(t+h) \geq 0$ . This implies that for  $h \leq S$  the inventory reduction shown in Lemma 2 is purely a reduction of the positive part of the inventory. The non-positive part of the inventory is not influenced. So, the inventory cost for the system without ADI is reduced by  $c(1 - \rho)h$ , compared to the system without ADI. This completes the proof.  $\square$

The last question is whether this order base-stock policy for the system with ADI that corresponds to the optimal base-stock policy for the system without ADI, is optimal. This is considered in Lemma 5.

**Lemma 5** *Let  $S$  be the base-stock level of the optimal base-stock policy for the system without ADI. Then the order base-stock policy for the system with ADI =  $h$  ( $h \leq S$ ) that corresponds to this optimal policy, is also optimal. The cost reduction effect of ADI is equal to  $c(1 - \rho)h$ .*

*Proof* Let  $S$  be the base-stock level of the optimal base-stock policy for the system without ADI and let  $R_o$  be the average reward. Corollary 2 shows immediately that application in the system with ADI ( $h \leq S$ ) of the order base-stock policy that corresponds to this optimal policy leads to a reduction of the cost of  $c(1 - \rho)h$  per unit of time. So, the average reward is equal to  $R_o + c(1 - \rho)h$ . What is left to prove is that this policy is optimal for the system with ADI. Suppose there is a better policy,  $p_o^h$ , for this system with ADI and let  $R_o^h$  be its average reward. So,  $R_o^h > R_o + c(1 - \rho)h$ . Consider the corresponding policy for the system without ADI,  $p_o^0$ . Let  $R_o^0$  be its average reward. Corollary 1 implies  $R_o^0 \geq R_o^h - c(1 - \rho)h$ . Together with  $R_o^h > R_o + c(1 - \rho)h$ , this implies  $R_o^0 > R_o$ . This is a contradiction and completes the proof.  $\square$

*Remark* The optimality of order base-stock policies and the corresponding inventory reduction only holds when  $h \leq S$ . This however is not a serious limitation in many cases. For instance, if the optimal base-stock level of the system without ADI is large, the critical threshold is just as large. This is a case where ADI can help in a significant manner. Inversely, if the optimal base-stock level without ADI is small, the critical threshold must be small but this may not be disturbing since this case corresponds to a system with low inventory related costs to begin with and the potential savings are limited regardless of ADI.

## 5 Extensions and limitations

This section explores whether the model assumptions of the previous section can be relaxed.

### Extensions

The lemmas show that the correspondence of the system with ADI and the system without ADI holds under a more general set of assumptions. The crucial equality is (see lemma 2):

$$I^h(t+h) = V(t) - L^h(t, t+h) \cdot 1.$$

This equality depends only on the way the capacity is modeled. The most important extensions are mentioned below.

It is clear from the analysis that the results also apply to the case with positive production throughput time  $\tau$ , as long as the horizon  $h > \tau$ . The situation without ADI has to be replaced by the situation with ADI equal to  $\tau$ . The effect of extra ADI is a reduction of the inventory cost with  $c(1-\rho)(h-\tau)$ . The case with  $h \leq \tau$  can be reduced to the case without ADI by using the modified inventory (= inventory plus work-in-process minus demand as far as it is known, Gallego and Ozer 2001).

The results also hold for the pure backlogging case (no rejection of customer orders) and the pure lost-sales case (no late deliveries because of proper order acceptance). This is immediately clear from inspection of the lemmas. Or, consider an alternative formulation of the problem in the pure backlogging case where the objective function is to minimize the average inventory holding cost subject to a time-average backorder-related service level constraint (such as the probability of stock-out or the expected number of backorders less than a specified value). The inventory cost reduction result with ADI naturally extends to this case. Note that the backorder levels of the systems with and without ADI are identical in distribution; therefore, regarding the service level constraint, both systems perform identically.

Most results also hold if production is not in units, but in larger batches. The difficulty is to prove that the production part of the optimal policy is of the base-stock type. That proof requires the invariance under the minimization operator of convexity properties of the value function. The proof is complicated by the existence of batches  $> 1$ , by the combination of production and order acceptance decisions and by the necessity to add the remaining length of the production run to the state space. Notice that the main result (the effect of ADI) holds also if the optimal policy for the system without ADI is not a pure base-stock policy but if there is some  $S > 0$ , such that production is started or continued if the inventory  $I \leq S$ .

The production varies between 0 and 1. The results remain valid if the production varies between some level  $p$ ,  $0 < p < 1$ , and 1. In that case, it may

also be expected that under general conditions the optimal policy for the system without ADI is of the critical level type: switch to the maximal production speed if the inventory is less than a certain critical level.

The Poisson assumption is also not really necessary to establish the relationship between corresponding policies for the systems with and without ADI. However, relaxing this assumption makes the structure of the optimal policies more complicated. For instance, the optimal policy for the system without ADI may not be a base-stock policy.

### Limitations of the approach

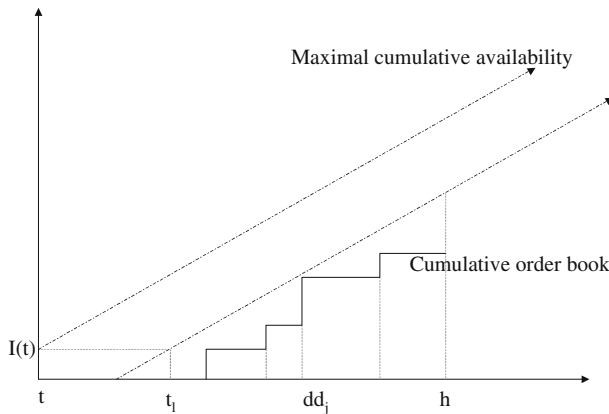
The assumptions with respect to the production are essential in the approach. Karaesmen et al. (2002) consider the case with a stochastic production time. They assume a geometrically distributed production time. Here, it would be more straightforward to assume a negative exponentially distributed production time. It is possible of course to define *corresponding* policies in this case as well. And the virtual inventory can be defined in the same way as in the continuous production case. But the relationship between the virtual inventory at time  $t$  and the actual inventory at time  $t + S$  (compare lemma 4) is less attractive. The production is only *in average* equal to the foreknowledge horizon and can be larger as well as smaller in this case. This makes it possible that  $V(t) = 0$  is followed by  $I(t + S) > 0$  or by  $I(t + S) < 0$ . The discontinuity at 0 destroys the clear relationship between  $V(t)$  and  $I(t + S)$ .

The same problems arise with more classes of customer orders, each with their own customer order lead time. A specific case deals with one class of orders that require immediate delivery and cannot be rejected and another class of orders with a fixed positive lead time. This case is equivalent to the case with unreliable production. This unreliability destroys the relationship between  $V(t)$  and  $I(t + S)$ .

## 6 Longer customer lead times

For longer horizons (when  $h > S$ ), the results of Sect. 4 do not hold any more. The results of Sect. 4 show straightforwardly that the effect of ADI for a horizon  $h > S$  is bounded from above by  $c(1 - \rho)h$ . That is because the inventory reduction is not fully concentrated any more on the stretches with positive inventory. But it has to be investigated how close we can get to this bound.

The problem with an order base-stock policy when  $h > S$  is that it may lead to a situation where the policy does not prescribe production because  $V(t) > S$ , while this production was presupposed earlier, when the system was confronted with a new order arrival. There are two straightforward heuristics for this case. The first is to use order base-stock policies, but to introduce a critical limit for  $h$  beyond which customer orders are not taken into account any more. See Karaesmen et al. (2002) for an application to the backlogging case. The other possibility is to use extra production checks. Let  $S$  be the base-stock level of the



**Fig. 2** Necessity of more detailed production checks

optimal (base-stock) policy for the system without ADI. If the order base-stock policy that corresponds to this optimal policy is used for the case with  $h > S$ , there is a certain risk that production is turned off, because the virtual inventory,  $V(t) > S$  which still leads to stock-outs. See Fig. 2 for an illustration. If the production start is postponed beyond time  $t_l$ , a stock-out results.

A stock-out is inevitable as soon as  $I(t) - C(t, dd_j) + 1(dd_j - t) < 0$ , with  $dd_j$  being the due date of order  $j$  and  $C(t, dd_j)$  the sum of the orders on the interval  $(t, dd_j]$ . Stock-outs can be prevented by adding the following checks:

$$I(t) - C(t, dd_j) + 1 \cdot (dd_j - t) \geq 0, \quad \forall \text{ orders } j \text{ in the order book.}$$

Policies satisfying these conditions are called *reliable*. They guarantee that if it is possible on arrival to deliver an order on time, the order is indeed delivered on time. The structure of the checks shows right away that they are not necessary indeed if  $h \leq S$ . Such reliable policies are applicable in a much wider range of  $h$ . Wijngaard (2004) shows that the reliable policy that results from the optimal base-stock policy gives good results for the pure backlogging case as well as the strict order acceptance case. Illustrative simulation results for a backlogging case with Poisson arrival, customer order size equal to 1, arrival rate 0.9, and  $c = 1, b(x) = 4x$  are given in Tables 1 and 2. Since  $c = 1$ , the upper bound for the cost reduction due to ADI is equal to  $(1 - \rho)h$ , with  $\rho$  the sample path utilization rate. The length of the simulation run is 100,000 time units. The optimal base-stock level for the system without ADI is  $S = 7$ . Since the customer order sizes are all equal to 1, it is optimal to accept maximally in the case without ADI: accept if  $I(t) \geq 1$ . The corresponding order base-stock policies in case of ADI accept an order if  $V(t) \geq 1$ . Table 1 gives the result for the pure order base-stock policies and Table 2 for the policies with the check added (reliable policies).

The results show that for a small  $h$  ( $h = 5$ ) the stock-out cost is not influenced by the ADI. The inventory cost reduction is equal to the total cost reduction and



**Table 1** Results for order base-stock policies

$h$	$S$	Inventory cost	Stock-out cost	Total cost	Total cost reduction	$(1 - \rho)h$
0	7	4.289	3.374	7.663	—	—
5	7	3.803	3.374	7.177	0.486	0.486
10	7	3.320	3.381	6.701	0.962	0.971
15	7	2.874	3.539	6.413	1.250	1.457
20	7	2.500	3.986	6.486	1.177	1.942
20	8	3.263	3.040	6.303	1.360	1.942
20	9	4.089	2.343	6.432	1.231	1.942

**Table 2** Results for reliable policies

$h$	$S$	Inventory cost	Stockout cost	Total cost	Total cost reduction	$(1 - \rho)h$
0	7	4.289	3.374	7.663	—	—
5	7	3.803	3.374	7.177	0.486	0.486
10	7	3.329	3.362	6.691	0.972	0.971
15	7	2.989	3.253	6.242	1.420	1.457
20	7	2.801	3.094	5.894	1.768	1.942
20	6	2.267	3.607	5.874	1.788	1.942
20	5	1.839	4.150	5.989	1.774	1.942

equal to the upper bound  $(1 - \rho)h$ . For larger  $h$ , the stock-out cost is increasing because of the above mentioned effect. Due to this, the rate at which the total cost reduction grows as function of  $h$  is decreasing for large  $h$ . By increasing the base-stock level it is possible to partly neutralize this effect. But the total cost reduction remains rather far from the bound  $((1 - \rho)h)$ . Table 2 gives the results for the policies with the check added (reliable policies).

Here the total cost reduction remains much closer to its upper bound. Due to the checks the inventory in the system becomes larger than in the case without checks. It is possible to compensate this by reducing the base-stock level. See the last lines in Table 2. This leads to even lower total cost.

## 7 Conclusions and suggestions for further research

This paper investigates the effect of ADI in a single item inventory model with constrained production capacity. Production is deterministic and continuous. So, the uncertainty comes only from the stochastic demand. The model combines the backlogging case and the lost-sales case. The decisions regard order acceptance and production. The ADI is modeled by assuming that the customer orders are known  $h$  time units in advance. The paper shows that as long as the foreknowledge horizon is smaller than a certain threshold value, the optimal policy is an order base-stock policy. This means that the orders in the order book may be aggregated over time. The threshold value is equal to the optimal base-stock for the system without ADI. The effect of ADI is equal to  $c(1 - \rho)hP$ ,

with  $P$  being the production speed. Recall that in the analysis we worked with  $P = 1$ , by changing the units of time. Here we go back to arbitrary units of time. The combination of backlogging and lost-sales leads directly to a new heuristic for the case with a larger  $h$ .

Future research has to go in three directions. In the first place it is useful to explore whether it is possible to broaden the result that it is allowed to aggregate the orders in the order book. The limitations of the approach applied here are indicated in the paper, but the result may be more general than the approach. An interesting case is the one with variable customer lead times.

The second direction is research about the applicability of the heuristic developed here. The idea of applying *reliable* heuristics is inspired by the order acceptance/lost-sales case, but turns out to be also useful in the backlogging case.

The third field for further research is the extension to more products. If the production runs may be small, the inventory positions can be kept close to each other and the aggregate inventory pattern gives a good indication for the item inventory patterns. This suggests that it may be possible to apply the results derived here for the single-item case in a multi-item case with small production runs.

## References

- Beckmann M (1961) An inventory model for arbitrary interval and quantity distribution of demand. *Manage Sci* 24:819–826
- Buzacott JA, Shanthikumar JG (1994) Safety stock versus safety time in MRP controlled production systems. *Manage Sci* 40:1678–1689
- Gallego G, Ozer O (2001) Integrating replenishment decisions with advance order information. *Manage Sci* 47:344–360
- Gayon JP, Benjaafar S, de Véricourt F (2004) Using imperfect demand information in production–inventory systems with multiple demand classes. Working Paper, Ecole Centrale, Paris
- Graves SC, Willems SP (2003) Supply chain design: safety stock placement and supply chain configuration. In: De Kok AG, Graves SC (eds) *Handbook in operations research and management science*, volume 11: supply chain management: design, coordination and operation. Elsevier, Amsterdam
- Gullu R (1995) Optimal production/inventory policies under forecasts and limited production capacity. Technical Report, METU, Ankara
- Gullu R (1996) On the value of information in dynamic production/inventory problems under forecast evolution. *Naval Rese Logistics* 43:289–303
- Hariharan R, Zipkin P (1995) Customer-order information, lead times and inventories. *Manage Sci* 41:1599–1607
- Inderfurth K (1991) Safety stock optimization in multi-stage inventory systems. *Int J Prod Econ* 24:103–113
- Karaesmen F, Buzacott JA, Dallery Y (2002) Integrating advance order information in production control. *IIE Trans* 34:649–662
- Karaesmen F, Liberopoulos G, Dallery Y (2003) Production/inventory control with advance demand information. In: Shantikumar JG, Yao DD, Zijm WHM (eds) *Stochastic modelling and optimization of manufacturing systems and supply chains*. International Series in Operations Research and Management Science 63, Kluwer Academic Publishers, Boston, pp 243–270
- Karaesmen F, Liberopoulos G, Dallery Y (2004) The value of advance demand information in production/inventory systems. *Ann Oper Res* 126:135–157

- Karlin S, Scarf HE (1958) Optimal inventory policy for the Arrow Harris Marchak dynamic model with a time lag. In: Arrow, Karlin and Scarf (eds.) *Studies in the mathematical theory of inventory and production*. Stanford University Press, Stanford
- Liberopoulos G, Koukournialos S (2005) Tradeoffs between base-stock levels, numbers of Kanbans, and production lead times in production/inventory systems with advance demand information. *Int J Produ Econ* 96:213–232
- Ozer O, Wei W (2004) Inventory control with limited capacity and advance demand information. *Oper Res* 52:988–1000
- Puterman ML (1994) *Markov decision processes*. Wiley, New York
- Ross SM (1970) *Applied probability models with optimization applications*. Holden-Day
- Simpson KF (1958) In-process inventories. *Oper Res* 6:863–873
- Toktay LB, Wein LM (2001) Analysis of a forecasting–production–inventory system with stationary demand. *Manage Sci* 47:1268–1281
- Wijngaard J (2004) The effect of foreknowledge of demand in case of a restricted capacity: the single-stage, single-product case. *Eur J Opera Res* 159:95–109
- Zipkin PH (2000) *Foundations of inventory management*, McGraw-Hill